DOCUMENT RESUME

ED 288 220 CS 505 798

AUTHOR Carey, Joyce; Rodier, Robin
TITLE Judging the Judges: A Content Analysis of Interpretation Ballots.
PUB DATE 5 Nov 87
NOTE 21p.; Paper presented at the Annual Meeting of the Speech Communication Association (73rd, Boston, MA, November 5-8, 1987).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitudes; Communication Research; *Competition; Content Analysis; Evaluation Criteria; *Judges; *Oral Interpretation; *Performance Factors; *Persuasive Discourse; Speech Communication; *Student Evaluation; Verbal Communication

ABSTRACT

To analyze the usefulness of forensics judges' written comments and to understand how to improve ballots, a content analysis of 175 ballots from schools in a competition district was conducted. Categories for analysis included the following: (1) quantity of comments; (2) format of the ballot; (3) types of comments (including selection, presentational skills, personal comments to contestants, judge disclosure of personal preferences or judging style, and comments written in the form of questions); (4) the value dimension of the comments; (5) the presence or absence of advice; (6) justification of rank; and (7) courtesy comments. The resulting data showed that judges wrote an average of 11 comments per ballot, using sentence format far more often than fragment or other style. An overwhelming majority of comments dealt with presentational skills, very few dealt with the choice or merit of the material. Personal comments to the contestants occurred with great frequency and were often very "personal" in nature. Overall, more negative comments were written than positive, although more than half the ballots contained a good balance of comments. An overwhelming number contained prescriptive advice to the student. However, judges rarely justified their ranking decisions; this is suggested as one area in which judges may best improve their ballot writing. (JG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED288220

Judging the Judges: A Content Analysis of Interpretation Ballots

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Joyce Carey

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

By Joyce Carey and Robin Rodier

Submitted to the Speech Communication Association
National Convention

Boston, Massachusetts

November 5, 1987

CS505798

BEST COPY AVAILABLE

Judging the Judges: A Content Analysis of Oral Interpretation Ballots

At the conclusion of a forensic tournament when ballots are returned to students there is usually an ensuing uproar from discontented students. Typical comments include, "Great! another ballot with no comments!" or "I can't believe this judge gave me glowing comments and a rank of 5," or "This ballot tells me nothing about what I can do to improve." or I got a good rank but I don't know why because everything is negative." As this panel sets out to "judge the judges", we are interested in looking at the ballots they write. Is the criticism they receive from students warranted? Are the ballots truly the educational and coaching tools they should be? By analyzing the two large divisions of forensics: public address and oral interpretation events, we will be able to identify the feedback judges are giving to students at forensic tournaments while preserving the nuances of each division. What are the nature of comments being written? What is the quantity of comments? What is the relationship of positive to negative comments on ballots? Is there a common format for writing comments? These are the questions which prompted this analysis of judges ballots in oral interpretation.

This type of study is important for several reasons. First, ballots serve an important role in the educational process. Corbin (1968, p. 91), explains, "evaluation is a matter of discriminating among the superior, the good, the mediocre and the inadequate. This discrimination serves two educational purposes: it helps establish standards and it provides a basis for improvement."

Furthermore, the ballot serves an important justificatory function in a competitive environment. Allen (1982, p. 15) wrote in an action caucus on oral interpretation, "One way or another, critics must be held accountable for their estimates of performance. They must be able to defend the rankings they assign, and this means being able to justify specific hierarchical standards which are applied to performance. Otherwise, we must accept that a great deal of judgement is, alas, based on impressions only, and that such impressions are based on conjecture and opinion as much as on fact."

Second, research in this area can supplement past studies in the area of performance criticism. Beverly Whitaker Long probably wrote the seminal work in this area. Her series of articles in The Speech Teacher built the argument for basing critical remarks on textual evidence (Long, 1969). Other scholars have examined such issues as reliability and validity of instrument measures. (Lee and Hewes, 1975). Seedorf (1950) and Lewis & Larsen, (1981) suggest there is a need for increased agreement between judges

although they disagree with each other whether judge training will increase that agreement.

More recent studies (Pearse and DeArmond, 1980) have given us insight into what type of comments students find most helpful. Their study found the following areas as most valuable feedback on student performance. Approximately 40% of respondents rated comments on technical areas of performance as most useful. Twenty-five percent of respondents thought important comments were those which commented about the understanding of speaker's stance in the performed text. Since we know what types of comments students find valuable, we can use it as the basis for analyzing what judges are actually writing on their ballots. A content analysis of the broad categories of forensics is needed as the first step in understanding how to improve ballots, in order to reduce the anger, frustration, and confusion students feel when reading them.

Method

One hundred and seventy-five ballots were collected and analyzed from the four categories of oral interpretation including prose (44 ballots), poetry (44 ballots), dramatic interpretation (44 ballots), and dramatic duo (43 ballots). Ballots were requested and received from schools throughout District IV. Of the 305 ballots received 175 were selected to ensure the widest variety of judges and contestants. The ballots were from 114 different

judges and 104 different students. The ballots used were all the same format and size (6x8). They were essentially blank, with no suggested criteria for evaluation. (Using a criteria-based ballot such as the one proposed at the Forensics Development Conference could influence the quantity and nature of comments.) To ensure some independence, no two contestants from the same round were chosen. The ballots came from tournaments which were nationally representative, including the American Forensic Association National Individual Events Tournament held in San Diego, California, as well as invitationals held in North Dakota, Minnesota, Wisconsin, and Texas. No novice tournament ballots were used.

The method of interpretation of data is referred to as content analysis. Berelson (1952, p. 18), defines content analysis as a research technique for the objective, systematic, and quantitative description of the manifest content of communication. In a later chapter entitled, "Qualitiative Content Analysis", Berelson (1952, p. 116) stresses that quantification need not be numerical but may take the form of words such as "more" or "always" or "increases." George (1959, p. 10), in Trends in Content Analysis, identifies this technique as "non-frequency" content analysis which uses the presence or absence of a certain content characteristic or syndrome as a content indicator in an inferential hypothesis." This study will analyze the ballot content according to 5 categories although no attempt will be made to assign numerical values to items. A

discussion of the method of selection and categorization should help clarify the technique.

Quantity of Comments

Trying to answer the question posed in the introduction about the quantity of comments students receive from judges, the first category simply counted the most basic comment units. The comments were broken down into the smallest unit possible with no judegments made about the quality of the comment. Thus, a generic comment like "very enjoyable", or "good job", was counted in order to identify the total number of comments being made. Similarly, if a comment was a restatement, it was still counted twice. For example, "the mother's voice was weak and although she is only 50 years old, you make her sound 80. This was reported as two comments although they say approximately the same thing. Thus, a picture emerges which does not emphasize quality or worth of the comment, but simply tries to identify the number of comments that exist on the ballot.

Format of Ballot

A second category of interest was the format of the ballot. Each ballot was judged by the general "look" of the ballot. Specific categorizations included: sentence, fragment, outline, and stream of consciousness ballot format. Sentence ballots were characterized by completeness of thought and presence of subject, verb and proper punctuation. Fragmented ballot styles were

typified by half sentences, use of key words and comments which usually appeared in list form. (Some even included symbols such as "+" and "*" to enumerate the ideas.) The outline format contained comments which were grouped and labeled by ideas such as, introduction, delivery, selection, etc. in a visual outline format. Stream of consciousness ballots gave the general appearance of fluidity, where ideas seemed to flow with little punctuation or separation.

Types of Comments

The third category had five subdivisions. We were looking for specific comments related to: 1) selection 2) presentational skills 3) personal comments to contestants 4) judge disclosure of personal preferences or judging style and, 5) comments written in the form of questions. The number of comments which fell into each of these five categories were counted. The only categories which were mutually exclusive were selection and presentational comments. Comments about selection included remarks about cutting and choice of material. Presentational skills were identified as those which focused on technique or performance including focal points, eye contact, rate, characterization, or specific line readings. Personal comments were those made directly to the contestant which usually commented about appearance, social life, etc. They included such comments as,"Nice perm, Amy!" "Great new suit, Dan." or "You are such a talented reader, Julie!" Another subset of this category coded comments which disclosed the

judge's individual preference or judging style. An example of this type of comment is, "I think you should consider more suggestion of the drinking rather than pantomiming it - this seems more within the interpretive realm." or "I don't like Shakespeare performed in the forensic settting." Finally, we counted the number of comments which took the form of questions. "Is this where the play really ends?" is one example of this type of "comment."

Valence

The fourth coding category identified the value dimension of comments. Comments were classified into five value dimensions: very positive, positive, neutral, negative, and very negative? We identified the "very positive" and "very negative" comments by the use of emphatic adjectives and punctuation. "The mother's voice was superb!" was coded as "very postive" while, "You're rocking back and forth, and it is extremely distracting!" was coded as a "very negative" comment. In contrast, simple positive comments included remarks like, "good introduction", and "nice pacing." Simple negative comments included, "Your cutting made the plot line hard to follow." or " You are speaking so quietly, I can't understand you." Comments which seem to be void of any positive or negative slant were coded as neutral. Special care was taken to ensure that comments were coded on face value and not by what they implied. This was important because some comments were written neutrally although they implied that the

student had done something "wrong". For example, "double check your focal points." or "remember to give both title and author in your introduction." All questions were also coded as neutral comments. All comments in this category of value dimension were mutually exclusive.

Advice

The fifth category analyzed comments for the presence or absence of prescriptive advice. If the ballot contained more than one piece of advice or prescription, the quantity was also noted. In order to be considered "prescriptive", the comment needed to extend beyond "what was wrong" to make some suggestion of how to correct it. "The focal points are too far apart," would not be considered a prescriptive comment, but a similar comment, "Bring your focal points closer together," would be prescriptive.

Justification of Rank

The sixth category looked for the presence or absence of comment or comments which justified the rank or rate given to the student. This justification had to be directly stated rather than indirectly implied. In other words, we were looking for specific statement(s) which told students, "you got this rank/rate because..." rather than accepting the implication that, "if you read my ballot you would know why you received that rank/rate."

Courtesy Comments

In the final category we were interested in the cliche comments like, "tough round,""thank you," and "overdone material". We were also looking to see if some indication was given on the ballot that the judge had timed the performance.

Results

Of the 175 ballots analyzed there were a total of 1877 comments, an average of 10.72 comments per ballot. The range of comments was 4-22.

Sixty-eight percent (n=119) of the ballots were written in sentence format while thirty percent (n=54) were classified as fragment style of ballot writing. The remaining two types were used infrequently. Only 1.14 percent of ballots (n=2) were considered stream of consciousness and there were no ballots classified as outline format.

In the second category, 77% (n=1456) of the comments dealt with the presentational aspect of a student's performance compared to only 7.51% (n=141) comments about their selection. In the other subdivisions of this category, 4.79% (n=90) of ballots contained information which was considered personal to the contestant. Comments which identified judges preferences or judging style occurred on 3.72% (n=70) of ballots. Eighty comments, or about 4.26%, were worded as questions on the ballots coded. There were also about 65 coded comment units which did not fit into any of the above categories.

When comments were coded according to a value dimension, 33.46% (n=628) were positive, 18.70% (n=351) were neutral, and 47.84% (n=898) were negative. The specific breakdown into the five areas was as follows: very positive 7.51% (n=141), positive 25.95% (n=487), neutral 18.70% (n=351), negative 47.52% (n=892), and very negative .32% (n=6). In addition to the overall relationship of positive to negative comments, we also calculated the percentage per ballot. Assuming that a "perfect" ballot would have a nearly equal number (40-60%) of positive and negative comments, we identified those ballots which did not have a proportion of positive to negative in that range. We found that 32.5% (n=57) contained more than 60% negative comments. Only 13.14% (n=23) were overly positive and in 54.28% of the cases (n=95) the comments were within the range suggesting they were balanced.

Of the 175 ballots analyzed, 87.42% (n=153) contained some prescriptive comments. There was a total of 475 comments coded as prescriptive or an average of just over three comments per ballot. Only 12.58% (n= 22) of the total ballots did not contain any prescriptive remarks.

Ninety-seven percent of ballots (n= 170) did not justify rank or rate on the ballot. Only 2.86% (n=5) gave any direct justification for the rank assigned.

When the presence or absence of courtesy comments was coded, 18.86% percent of ballots (n=33) contained a "thank you" or similar type of remark for the performer. Six ballots (3.42%)

told the performer it was a "tough round." No ballots made any reference to the literature being overdone. Only 46% (n=81) of judges gave any indication on the ballot that they had timed the performance.

Discussion

The data from this study has given us some interesting insights into the content of ballots. With an average of nearly eleven comments per ballot it is apparent that judges are writing an adequate number of comments on student ballots, although we cannot draw any inferences about the quality of those comments.

Not surprisingly, the sentence format was used far more often than any other ballot style. This may be because of the ease in writing such a ballot as well as the ease in understanding its content.

When analyzing the content of the comments several interesting observations can be made. First, an overwhelming majority of comments on the ballot dealt with presentational skills versus the material itself. In fact, judges wrote very little about choice of material. The sensitive issue of overdone was not mentioned on any of the 175 ballots. Either judges are no longer concerned about this issue or student's are not choosing to perform overdone material. Another corollary of this category was the issue of literary merit. Of the 141 comments on selection, only four mentioned literary merit on the ballot.

The prevalence of personal comments on forensic ballots is also an interesting phenomenon. They occurred with great frequency on student's ballots, and were often very "personal" in nature. This may be a distinctive feature of forensics in District IV or it may simply be a type of judging style. It is our opinion that the District from which this sample was drawn may have influenced the data. District IV is very close-knit and contains many coaches, judges, and students which see each other at tournaments on a weekly basis. This fosters a high degree of friendliness, and may explain the frequency of personal comments. Whether or not this is appropriate or useful to include on the ballot is a matter of personal style but this study shows it is certainly a common occurrence. One judge went so far as to note on the ballot, "Good to see you. I can't believe you only placed third last weekend at Mankato." It can be said that ballots examined were quite personalized rather than being cold and objective evaluative statements.

In addition to personalizing the ballot to contestants, judges are also using the ballot to disclose their own personal feelings and opinions about what they are hearing. A number of ballots expressed judges likes and dislikes. Judges most commonly disclosed their personal opinion about the distinction between acting and interpretation. Other themes which emerged in this category included comments on how material was cut, the inclusion of offensive language, and personal preference on how specific lines should be read. While the majority of comments were written

in declarative form, questions were also prevalent on student ballots. Some of these were truly questions which seek clarification such as,"How old is the little boy in this selection?" but others were statements which were simply transformed into questions. For example, "Don't you think there should be more tension between the mother and daughter as they talk about the father? "This is probably simply a judging style but it tends to avoid directness in the comment. The rhetorical question was also used in these ballots. "Don't gesture with your book, Okay?" Use of questions may simply be a type of judging style or they may fulfill the need to seek clarification at the same time witholding judgment on the issue. I don't think the judge ever expects to get an answer, but it encourages the student to ask questions of their own performance.

Identifying the positive or negative nature of comments was probably the most intriguing aspect of this study. It pointed out that overall more negative comments are written on ballots than positive comments. Of the 1877 total comments analyzed, there were almost twice as many negative comments as positive comments on the ballots. This was somewhat deceiving however, because the total number did not reflect the proportion of positive to negative comments which existed on the individual ballot. When individual ballots were analyzed for a balance between positive and negative comments, we found that over half (54%) of the ballots contained a good balance of comments. However, those ballots which were unbalanced tended to be overly negative rather

than overly positive. It would be interesting to correlate those ballots which were unbalanced with the ranking received.

The contrast between use of very positive and very negative comments was significant. Very positive comments occurred almost 25 times more often than very negative comments. This may indicate the judge's willingness to be very positive about student performances but a reluctance to be very negative. After all, there is a difference between constructive criticism and very discouraging, disheartening remarks. If a judge found fault with a performance it could be expressed in such a way that improvement was called for but not written in such a way that the student would be devastated. In the few instances where we coded very negative comments this was probably the most blunt. "This is ridiculous, I have seen you in every interp event at this tournament except duo, and they all sound the same! You are not stretching yourself but relying on what works!"

Some judges rely heavily on neutral comments which are typically very noncommittal in nature. A closer look at these neutral comments revealed an underlying negative idea that was expressed in a very neutral manner. For example, "If you look back at the whole play, it may change your approach to the son's character." Although the judge does not say that the "appproach" is currently bad, it implies that it needs to be reexamined. One encouraging conclusion supported by this study was the overwhelming number of ballots which contained prescriptive advice to the student. Some were very specific to line readings

while others were more general presentational suggestions. In order to be considered prescriptive, however, the comment had to suggest what was wrong as well as offer suggestions for improving the weakness. Only 22 of the 175 ballots did not offer ways to improve the performance. In fact, there was an average of just over 3 prescriptive comments per ballot. This does suggest that judges are not just making negative comments but they are also offering students ideas to improve their future performances.

While ballots do contain prescriptive information this study showed that only rarely do judges explicitly justify their decision on the ballot. In this category we were looking for comments which would give students a reason for their rank or rating. This type of comment appeared on only 5 ballots. Most of these comments either claimed that "being overtime" affected their rank or stated a comparison to others in the round. One judge commented, "Your material was not as difficult as others in this round. You did a fine pike, but they did a half gainer." It is possible that judges believe they are justifying their ballots indirectly. In other words, if you read my ballot you would know what you did wrong, and thus understand why you received that rank/rate. This may be the rationale for such overwhelmingly negative comments on ballots. This phenomenon may be one of the reasons students get frustrated when looking at the ballots. There is often no clear logical or apparent reason for the rank or rate. If the ballot is filled with negative comments then they may deduce that they got the "1" because they did fewer things

wrong than the other contestants. Similarly, if the ballot is filled with positive comments and they got a 5 they have no clear explanation for the judges decision. While it is impossible to understand justify this behavior, it is important to note that it is very uncommon for judges to justify their decison on the ballot.

A final bit of information emerged from this study regarding judges timing of performances. In this study, less than half of the judges timed, or at least gave any written indication of such on the ballot. Does this mean that judges do not consider time a judging criteria?

This study pointed out several interesting aspects of judges comments on interpretation ballots while it also made suggestions for future research in this area. Judges do seem to be doing a good job of writing an adequate quantity of comments, balancing them between negative and positive dimensions, and offering prescriptive advice. We would suggest that judges may best improve their ballots by writing more comments that justify their rank and rating decisions. Perhaps judges will always be the scapegoat for forensic students and the quality of their ballots has little to do with the criticism they receive, but it is our hope that this study may have made judges more aware of the nature and importance of the comments they write. After all, ballots play an important role in the educational process of literature performance and in the forensic activity. Beverly Whitaker Long, (1969, p. 193) may have said it best, "A reader

works with an art object, and his performance, if he actualizes that object, is also an art object. Achieved, his efforts have been considerable. Can the critic afford to be less careful in appraising that performance?"

Bibliography

Bacon, W. (1979). The art of interpretation. (3rd Ed.). New York: Holt Rinehart and Winston.

Berelson, B. (1952). Content analysis in communication research Glencoe, IL: Free Press.

George, A. (1959). Trends in content analysis, Urbana, IL: University oif Illinois Press.

Grimes, W.H. (1958). Oral interpretation and criticism: A bibliography. Western Journal of Speech Communication, 22 69-74.

Holloway, H. (November, 1982). Report on action caucus on oral interpretation in forensics A paper presented at the annual meeting of the Speech Communication Association. Louisville, KY.

Irwin, R. (1966). The bases of criticism in oral interpretation. Communication Quarterly, 14, 5-6.

Johnson, G. (1940). Studies in the art of interpretation. New York: Appleton-Century.

Lewis, J., and Larsen J. (1981). Inter-rater judge agreement in forensic competition. Journal of the American Forensic Association. 18, 9-16.

Lee, C., and Gura, T. (1982). Oral Interpretation. (6th Ed.). Boston, MA: Houghton Mifflin.

Lee, R., and Hewes, D. (1975). Reliability and validity of two empirical measures of oral interpretation. Journal of the American Forensic Association. 12, 61-69.

Long, B. (1969). Critical reasons and literature in performance. Speech Teacher, 18, 191-193.

Pearse, J., and DeArmond W. (1980). Critical judgements of performed texts: A survey. Communication Qaurterly, 28, 57-63.

Pelias, R. (1984). Evaluating interpretation events on the forensic circuit. Journal of the American Forensic Association, 20, 191-230.

Seedorf, E. (1950). Evaluating oral interpretation. Western Speech Journal, 14,28-30.